



OIKE

ENTERED

## RAW SEQUENCE LISTING

DATE: 07/02/2002

PATENT APPLICATION: US/10/078,650

TIME: 14:23:13

Input Set : A:\06501-101001.txt

Output Set: N:\CRF3\07022002\J078650.raw

4 <110> APPLICANT: Fujimoto, Katsumi  
5 Shin, Mei  
6 Kato, Yukio  
9 <120> TITLE OF INVENTION: NOVEL bHLH TYPE TRANSCRIPTION FACTOR  
10 GENES DEC2  
12 <130> FILE REFERENCE: 06501-101001  
14 <140> CURRENT APPLICATION NUMBER: US 10/078,650  
15 <141> CURRENT FILING DATE: 2002-02-19  
17 <150> PRIOR APPLICATION NUMBER: PCT/JP00/03991  
18 <151> PRIOR FILING DATE: 2000-06-19  
20 <150> PRIOR APPLICATION NUMBER: JP 11-233286  
21 <151> PRIOR FILING DATE: 1999-08-19  
23 <160> NUMBER OF SEQ ID NOS: 18  
25 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
27 <210> SEQ ID NO: 1  
28 <211> LENGTH: 3641  
29 <212> TYPE: DNA  
30 <213> ORGANISM: Homo sapiens  
32 <220> FEATURE:  
33 <221> NAME/KEY: CDS  
34 <222> LOCATION: (135)...(1580)  
36 <400> SEQUENCE: 1  
37 ctgcactgaa gagggagagc gagagagaga ctggagaagc acagatcccc ccaaggtctc 60  
38 ccaagcctac cgtccacag attattgtac agagcccaaa aaatcgaaac agaggaaacg 120  
39 aacagcagtt gaac atg gac gaa gga att cct cat ttg caa gag aga cag 170  
40 Met Asp Glu Gly Ile Pro His Leu Gln Glu Arg Gln  
41 1 5 10  
43 tta ctg gaa cat aga gat ttt ata gga ctg gac tat tcc tct ttg tat 218  
44 Leu Leu Glu His Arg Asp Phe Ile Gly Leu Asp Tyr Ser Ser Leu Tyr  
45 15 20 25  
47 atg tgt aaa ccc aaa agg agc atg aaa cga gac gac acc aag gat acc 266  
48 Met Cys Lys Pro Lys Arg Ser Met Lys Arg Asp Asp Thr Lys Asp Thr  
49 30 35 40  
51 tac aaa tta ccg cac aga tta ata gaa aag aaa aga aga gac cga att 314  
52 Tyr Lys Leu Pro His Arg Leu Ile Glu Lys Lys Arg Arg Asp Arg Ile  
53 45 50 55 60  
55 aat gaa tgc att gct cag ctg aaa gat tta ctg cct gaa cat ctg aaa 362  
56 Asn Glu Cys Ile Ala Gln Leu Lys Asp Leu Leu Pro Glu His Leu Lys  
57 65 70 75  
59 ttg aca act ctg gga cat ctg gag aaa gct gta gtc ttg gaa tta act 410  
60 Leu Thr Thr Leu Gly His Leu Glu Lys Ala Val Val Leu Glu Leu Thr  
61 80 85 90  
63 ttg aaa cac tta aaa gct tta acc gcc tta acc gag caa cag cat cag 458

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/078,650

DATE: 07/02/2002

TIME: 14:23:13

Input Set : A:\06501-101001.txt

Output Set: N:\CRF3\07022002\J078650.raw

64	Leu	Lys	His	Leu	Lys	Ala	Leu	Thr	Ala	Leu	Thr	Glu	Gln	Gln	His	Gln	
65			95				100						105				
67	aag	ata	att	gct	tta	cag	aat	ggg	gag	cga	tct	ctg	aaa	tcg	ccc	att	506
68	Lys	Ile	Ile	Ala	Leu	Gln	Asn	Gly	Glu	Arg	Ser	Leu	Lys	Ser	Pro	Ile	
69		110				115					120						
71	cag	tcc	gac	ttg	gat	gcg	ttc	cac	tcg	gga	ttt	caa	aca	tgc	gcc	aaa	554
72	Gln	Ser	Asp	Leu	Asp	Ala	Phe	His	Ser	Gly	Phe	Gln	Thr	Cys	Ala	Lys	
73	125					130					135					140	
75	gaa	gtc	ttg	caa	tac	ctc	tcc	cgg	ttt	gag	agc	tgg	aca	ccc	agg	gag	602
76	Glu	Val	Leu	Gln	Tyr	Leu	Ser	Arg	Phe	Glu	Ser	Trp	Thr	Pro	Arg	Glu	
77					145					150					155		
79	ccg	cgg	tgt	gtc	cag	ctg	atc	aac	cac	ttg	cac	gcc	gtg	gcc	acc	cag	650
80	Pro	Arg	Cys	Val	Gln	Leu	Ile	Asn	His	Leu	His	Ala	Val	Ala	Thr	Gln	
81				160					165					170			
83	ttc	ttg	ccc	acc	ccg	cag	ctg	ttg	act	caa	cag	gtc	cct	ctg	agc	aaa	698
84	Phe	Leu	Pro	Thr	Pro	Gln	Leu	Leu	Thr	Gln	Gln	Val	Pro	Leu	Ser	Lys	
85			175					180					185				
87	ggc	acc	ggc	gct	ccc	tcg	gcc	gcc	ggg	tcc	gcg	gcc	gcc	ccc	tgc	ctg	746
88	Gly	Thr	Gly	Ala	Pro	Ser	Ala	Ala	Gly	Ser	Ala	Ala	Ala	Pro	Cys	Leu	
89		190					195					200					
91	gag	cgc	gcg	ggg	cag	aag	ctg	gag	ccc	ctc	gcc	tac	tgc	gtg	ccc	gtc	794
92	Glu	Arg	Ala	Gly	Gln	Lys	Leu	Glu	Pro	Leu	Ala	Tyr	Cys	Val	Pro	Val	
93	205					210					215				220		
95	atc	cag	cgg	act	cag	ccc	agc	gcc	gag	ctc	gcc	gcc	gag	aac	gac	acg	842
96	Ile	Gln	Arg	Thr	Gln	Pro	Ser	Ala	Glu	Leu	Ala	Ala	Glu	Asn	Asp	Thr	
97				225					230					235			
99	gac	acc	gac	agc	ggc	tac	ggc	ggc	gaa	gcc	gag	gcc	cgg	ccg	gac	cgc	890
100	Asp	Thr	Asp	Ser	Gly	Tyr	Gly	Gly	Glu	Ala	Glu	Ala	Arg	Pro	Asp	Arg	
101				240					245					250			
103	gag	aaa	ggc	aaa	ggc	gcg	ggg	gcg	agc	cgc	gtc	acc	atc	aag	cag	gag	938
104	Glu	Lys	Gly	Lys	Gly	Ala	Gly	Ala	Ser	Arg	Val	Thr	Ile	Lys	Gln	Glu	
105			255					260					265				
107	cct	ccc	ggg	gag	gac	tcg	ccg	gcg	ccc	aag	agg	atg	aag	ctg	gat	tcc	986
108	Pro	Pro	Gly	Glu	Asp	Ser	Pro	Ala	Pro	Lys	Arg	Met	Lys	Leu	Asp	Ser	
109			270					275					280				
111	cgc	ggc	ggc	ggc	agc	ggc	ggc	ggc	ccg	ggg	ggc	ggc	gcg	gcg	gcg	gcg	1034
112	Arg	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Pro	Gly	Gly	Gly	Ala	Ala	Ala	Ala	
113	285					290					295					300	
115	gca	gcc	gcg	ctt	ctg	ggg	ccc	gac	cct	gcc	gcc	gcg	gcc	gcg	ctg	ctg	1082
116	Ala	Ala	Ala	Leu	Leu	Gly	Pro	Asp	Pro	Ala	Ala	Ala	Ala	Ala	Leu	Leu	
117				305						310					315		
119	aga	ccc	gac	gcc	gcc	ctg	ctc	agc	tcg	ctg	gtg	gcg	ttc	ggc	gga	ggc	1130
120	Arg	Pro	Asp	Ala	Ala	Leu	Leu	Ser	Ser	Leu	Val	Ala	Phe	Gly	Gly	Gly	
121				320					325					330			
123	gga	ggc	gcg	ccc	ttc	ccg	cag	ccc	gcg	gcc	gcc	gcg	gcc	ccc	ttc	tgc	1178
124	Gly	Gly	Ala	Pro	Phe	Pro	Gln	Pro	Ala	Ala	Ala	Ala	Ala	Pro	Phe	Cys	
125			335					340					345				
127	ctg	ccc	ttc	tgc	ttc	ctc	tcg	cct	tct	gca	gct	gcc	gcc	tac	gtg	cag	1226
128	Leu	Pro	Phe	Cys	Phe	Leu	Ser	Pro	Ser	Ala	Ala	Ala	Ala	Tyr	Val	Gln	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/078,650

DATE: 07/02/2002

TIME: 14:23:13

Input Set : A:\06501-101001.txt

Output Set: N:\CRF3\07022002\J078650.raw

```
129      350      355      360
131 ccc ttc ctg gac aag agc ggc ctg gag aag tat ctg tac ccg gcg gcg 1274
132 Pro Phe Leu Asp Lys Ser Gly Leu Glu Lys Tyr Leu Tyr Pro Ala Ala
133 365      370      375      380
135 gct gcc gcc ccg ttc ccg ctg cta tac ccc ggc atc ccc gcc ccg gcg 1322
136 Ala Ala Ala Pro Phe Pro Leu Leu Tyr Pro Gly Ile Pro Ala Pro Ala
137      385      390      395
139 gca gcc gcg gca gcc gcc gcc gcc gct gcc gcc gcc gcc gcc gcc gcg ttc 1370
140 Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Phe
141      400      405      410
143 ccc tgc ctg tcc tcg gtg ttg tcg ccc cct ccc gag aag gcg ggc gcc 1418
144 Pro Cys Leu Ser Ser Val Leu Ser Pro Pro Pro Glu Lys Ala Gly Ala
145      415      420      425
147 gcc gcc gcg acc ctc ctg ccg cac gag gtg gcg ccc ctt ggg gcg ccg 1466
148 Ala Ala Ala Thr Leu Leu Pro His Glu Val Ala Pro Leu Gly Ala Pro
149      430      435      440
151 cac ccc cag cac ccg cac ggc cgc acc cac ctg ccc ttc gcc ggg ccc 1514
152 His Pro Gln His Pro His Gly Arg Thr His Leu Pro Phe Ala Gly Pro
153 445      450      455      460
155 cgc gag ccg ggg aac ccg gag agc tct gct cag gaa gat ccc tcg cag 1562
156 Arg Glu Pro Gly Asn Pro Glu Ser Ser Ala Gln Glu Asp Pro Ser Gln
157      465      470      475
159 cca gga aag gaa gct ccc tgaatccttg cgtcccgaag gacggaggtt 1610
160 Pro Gly Lys Glu Ala Pro
161      480
163 caagcagagt gagaagttaa aataccctta aggaggttca agcagagtga gaagttaaaa 1670
164 tacccttaag gtctttaagg gaggaagtgt aatagatgca cgacaggcat aaacaagaac 1730
165 aacaaaacag gtgttatgtg tacattcgga gttcctgttt tgctcatccc gcaccacccc 1790
166 accctccaca cactaacatc cctttcttcc ccccaccagc tgtaaaagat cctatgcgaa 1850
167 agacactggc tctttttttt aatcccccaa ataaattttg ccccttttta ggccatgttc 1910
168 cattatctct taaaattgga acctaattcg agaggaagta agaagggtct gttctgtggc 1970
169 tgagctaggt gaaccccggy gtaggggaga gatgtaaca ctttgacgt ctttgaggtt 2030
170 gacatggaac agcaggtagt gtattatgag agctagttct caaagctgcc ctgcctgttt 2090
171 taggaggcgt tccacaaaca gattgaggct ctttttagaa ttgaatttac tcttcagtat 2150
172 tttctaattg tcagctttct aaaaggcata tatttttcaa agaagtgagg atgcagtttc 2210
173 tcacgttgca acctattctg aagtgggtta aatggatatc cttagtaact tgcactcgtt 2270
174 aaagaaacac ggagctgggc catcgtcaga actaagtcag ggaaggagat ggatgagaag 2330
175 gccagaatca ttctagtagc atttgctaac actttattga gaaattgacc atgaattaat 2390
176 ggactcatct taattttctc taagtccata tatagataga tatctatctg tacagatttc 2450
177 tatttatcca tagataggta tctatacata cacatctcaa gtgcactctat tcccactctc 2510
178 attaatccat catgttccta aatttttgta atcttactgt aaaaaaaagt gcactgaact 2570
179 tcaaaacaaa acaaaaaaca acaacaacaa aaaacaagtc caaactgata taccctatat 2630
180 tctgttaaaa ttcaaaagtg aacgaaagca tttaactggc cagttttgat tgcaaatgct 2690
181 gtaaagatat agaatgaagt cctgtgaggc cttcctatct ccaagtctat gtattttctg 2750
182 gagaccaaac cagataccag ataatcaca agaaagcttt tttaataagg cttaaaccaa 2810
183 gacctgtgct agatattttt agtttggtgc caaggtagca ctgtgagaaa tctcacttgg 2870
184 atgttatgta aggggtgaga cacaacagtc tgactatgag tgaggaaaat atctgggtct 2930
185 tttcgctcagt ttggtgcatt tgctgctgct gttgctactg tttgcctcaa acgctgtgtt 2990
186 taaacaacgt taaactctta gcctacaagg tggctcttat gtacatagtt gttaatacat 3050
```

## RAW SEQUENCE LISTING

DATE: 07/02/2002

PATENT APPLICATION: US/10/078,650

TIME: 14:23:13

Input Set : A:\06501-101001.txt

Output Set: N:\CRF3\07022002\J078650.raw

```
187 ccaattaatg atgtctgaca tgctattttt gtagggagaa aatatgtgct aatgatattt 3110
188 tgagttaaaa tatcttttgg ggaggatttg ctgaaaagtt gcacttttgt tacaatgctt 3170
189 atgcttggtta caagcttatg ctgtcttaaa ttatttttaa aaaattaaat actgtctgtg 3230
190 agaaaccagc tggtttagaa aagtttagta tgtgacgata aactagaaat tacctttata 3290
191 ttctagtatt ttcagcactc cataaattct attacctaaa tattgccaca ctattttgtg 3350
192 atttaaaaaat tcttactaag gaataaaaac tttaatatata gatatgatata tgtctaataa 3410
193 ttaaaaaaga cataatggat gctcaattag ttttaagata tctataacta tagggataca 3470
194 aatcactaca gttctcagat ttacaccttt tttttgtcat tggcttgatg tcacacattt 3530
195 ccaatctctt gcaagcctcc aggtctctgc tttgtctacc tgctcgttcc caatgtatct 3590
196 taatgaaaag tgcaaaaagaa aaacctacca attaaaaaaa aaaaaaaaaa a 3641
198 <210> SEQ ID NO: 2
199 <211> LENGTH: 482
200 <212> TYPE: PRT
201 <213> ORGANISM: Homo sapiens
203 <400> SEQUENCE: 2
204 Met Asp Glu Gly Ile Pro His Leu Gln Glu Arg Gln Leu Leu Glu His
205 1 5 10 15
206 Arg Asp Phe Ile Gly Leu Asp Tyr Ser Ser Leu Tyr Met Cys Lys Pro
207 20 25 30
208 Lys Arg Ser Met Lys Arg Asp Asp Thr Lys Asp Thr Tyr Lys Leu Pro
209 35 40 45
210 His Arg Leu Ile Glu Lys Lys Arg Arg Asp Arg Ile Asn Glu Cys Ile
211 50 55 60
212 Ala Gln Leu Lys Asp Leu Leu Pro Glu His Leu Lys Leu Thr Thr Leu
213 65 70 75 80
214 Gly His Leu Glu Lys Ala Val Val Leu Glu Leu Thr Leu Lys His Leu
215 85 90 95
216 Lys Ala Leu Thr Ala Leu Thr Glu Gln Gln His Gln Lys Ile Ile Ala
217 100 105 110
218 Leu Gln Asn Gly Glu Arg Ser Leu Lys Ser Pro Ile Gln Ser Asp Leu
219 115 120 125
220 Asp Ala Phe His Ser Gly Phe Gln Thr Cys Ala Lys Glu Val Leu Gln
221 130 135 140
222 Tyr Leu Ser Arg Phe Glu Ser Trp Thr Pro Arg Glu Pro Arg Cys Val
223 145 150 155 160
224 Gln Leu Ile Asn His Leu His Ala Val Ala Thr Gln Phe Leu Pro Thr
225 165 170 175
226 Pro Gln Leu Leu Thr Gln Gln Val Pro Leu Ser Lys Gly Thr Gly Ala
227 180 185 190
228 Pro Ser Ala Ala Gly Ser Ala Ala Ala Pro Cys Leu Glu Arg Ala Gly
229 195 200 205
230 Gln Lys Leu Glu Pro Leu Ala Tyr Cys Val Pro Val Ile Gln Arg Thr
231 210 215 220
232 Gln Pro Ser Ala Glu Leu Ala Ala Glu Asn Asp Thr Asp Thr Asp Ser
233 225 230 235 240
234 Gly Tyr Gly Gly Glu Ala Glu Ala Arg Pro Asp Arg Glu Lys Gly Lys
235 245 250 255
236 Gly Ala Gly Ala Ser Arg Val Thr Ile Lys Gln Glu Pro Pro Gly Glu
237 260 265 270
```

## RAW SEQUENCE LISTING

DATE: 07/02/2002

PATENT APPLICATION: US/10/078,650

TIME: 14:23:13

Input Set : A:\06501-101001.txt

Output Set: N:\CRF3\07022002\J078650.raw

238 Asp Ser Pro Ala Pro Lys Arg Met Lys Leu Asp Ser Arg Gly Gly Gly  
239           275                   280                   285  
240 Ser Gly Gly Gly Pro Gly Gly Gly Ala Ala Ala Ala Ala Ala Leu  
241           290                   295                   300  
242 Leu Gly Pro Asp Pro Ala Ala Ala Ala Ala Leu Leu Arg Pro Asp Ala  
243 305                   310                   315                   320  
244 Ala Leu Leu Ser Ser Leu Val Ala Phe Gly Gly Gly Gly Ala Pro  
245                   325                   330                   335  
246 Phe Pro Gln Pro Ala Ala Ala Ala Ala Pro Phe Cys Leu Pro Phe Cys  
247                   340                   345                   350  
248 Phe Leu Ser Pro Ser Ala Ala Ala Ala Tyr Val Gln Pro Phe Leu Asp  
249                   355                   360                   365  
250 Lys Ser Gly Leu Glu Lys Tyr Leu Tyr Pro Ala Ala Ala Ala Ala Pro  
251           370                   375                   380  
252 Phe Pro Leu Leu Tyr Pro Gly Ile Pro Ala Pro Ala Ala Ala Ala Ala  
253 385                   390                   395                   400  
254 Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Phe Pro Cys Leu Ser  
255                   405                   410                   415  
256 Ser Val Leu Ser Pro Pro Pro Glu Lys Ala Gly Ala Ala Ala Ala Thr  
257                   420                   425                   430  
258 Leu Leu Pro His Glu Val Ala Pro Leu Gly Ala Pro His Pro Gln His  
259                   435                   440                   445  
260 Pro His Gly Arg Thr His Leu Pro Phe Ala Gly Pro Arg Glu Pro Gly  
261           450                   455                   460  
262 Asn Pro Glu Ser Ser Ala Gln Glu Asp Pro Ser Gln Pro Gly Lys Glu  
263 465                   470                   475                   480  
264 Ala Pro

267 &lt;210&gt; SEQ ID NO: 3

268 &lt;211&gt; LENGTH: 32

269 &lt;212&gt; TYPE: DNA

270 &lt;213&gt; ORGANISM: Artificial Sequence

272 &lt;220&gt; FEATURE:

273 &lt;223&gt; OTHER INFORMATION: Artificially Synthesized Primer Sequence

276 &lt;400&gt; SEQUENCE: 3

277 ctattcgatg atgaagatac cccaccaaac cc

32

279 &lt;210&gt; SEQ ID NO: 4

280 &lt;211&gt; LENGTH: 25

281 &lt;212&gt; TYPE: DNA

282 &lt;213&gt; ORGANISM: Artificial Sequence

284 &lt;220&gt; FEATURE:

285 &lt;223&gt; OTHER INFORMATION: Artificially Synthesized Primer Sequence

288 &lt;400&gt; SEQUENCE: 4

289 gcaagtgggt gatcagctgg acaca

25

291 &lt;210&gt; SEQ ID NO: 5

292 &lt;211&gt; LENGTH: 21

293 &lt;212&gt; TYPE: DNA

294 &lt;213&gt; ORGANISM: Artificial Sequence

296 &lt;220&gt; FEATURE:

297 &lt;223&gt; OTHER INFORMATION: Artificially Synthesized Primer Sequence

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/078,650

DATE: 07/02/2002

TIME: 14:23:14

Input Set : A:\06501-101001.txt

Output Set: N:\CRF3\07022002\J078650.raw